Future of technology in NERC data models and informatics: outputs from InformaTEC

Andrew Kingdon[1], Jeremy R. A. Giles[1] and Jonathan P. Lowndes[2]

Author Affiliations

1. British Geological Survey, Keyworth, Nottingham, NG 12 5GG, UK
2. British Geological Survey, Murchison House, West Mains Road, Edinburgh EH9 3LA, UK

Corresponding author Andrew Kingdon (aki@bgs.ac.uk)

Abstract

The 'Big Data' paradigm will revolutionize understanding of the natural environment. New technologies are revolutionizing our ability to measure, model, understand and make robust, evidence-based predictions at increasingly spatial and temporal resolutions. Realising this potential will require reengineering of environmental sciences in the observation infrastructure, in data management and processing, and in the culture of environmental sciences. Collectively these will deliver vibrant, integrated research communities. Manipulating such enormous data streams requires a new data infrastructure underpinned by four technologies. Pervasive environmental sensor networks will continuously measure suites of environmental parameters and transmit these wirelessly to scientists, regulators and modellers in real time. Integrated environmental modelling will process data, streamed from sensor networks, using components synthesizing natural systems developed by domain experts, each of which will be linked at runtime to other expert developed components. Semantic interoperability will facilitate cross-disciplinary working, as has already happened within the biosciences so that data items can be exchanged with unambiguous, shared meaning. Cloud computing will revolutionize data processing allowing scalable computing close to observations on an as-needed basis. Leveraging the full potential of these technologies requires a major culture change in the environmental sciences where national and continental scale observatories of sensors networks become basic scientific tools.